Information-Theoretic Image Fusion Assessment Without Reference

Bruno Aiazzi
Italian National Research Council
225 PUBLICATIONS 3,910 CITATIONS

Stefano Baronti
Italian National Research Council
273 PUBLICATIONS 5,247 CITATIONS

Andrea Garzelli
Università degli Studi di Siena
225 PUBLICATIONS 3,575 CITATIONS

Filippo Nencini
Università degli Studi di Siena
52 PUBLICATIONS 1,749 CITATIONS

INFORMATION-THEORETIC IMAGE FUSION ASSESSMENT WITHOUT REFERENCE

Luciano Alparone(1), **Bruno Aiazzi**(2), **Stefano Baronti**(2), **Andrea Garzelli**(3), **Filippo Nencini**(3)

(1)*DET-UniFI: Department of Electronics and Telecommunications, University of Firenze, 3 Via Di Santa Marta, 50139 Firenze (Italy), Alparone@lci.det.unifi.it*

(2)*IFAC–CNR: Institute of Applied Physics "Nello Carrara", Italian National Research Council, 10 Via Madonna Del Piano, 50019 Sesto F.no (FI) (Italy), {B.Aiazzi S.Baronti }@ifac.cnr.it*

(3)*DII-UniSI: Department of Information Engineering, University of Siena, 56 Via Roma, 53100 Siena (Italy), {Garzelli Nencini}@dii.unisi.it*

ABSTRACT

In this work we investigate the use of Shannon's information theory for the goal of devising quality scores of image fusion results, that do not require reference originals. In particular, the mutual information between re-sampled original and fused MS bands is used to measure the spectral quality, while the mutual information between the Pan image and the fused bands yields a measure of spatial quality. The rationale is that the mutual information calculated either between any couple of bands, or between each band and the Pan image, should be unchanged after fusion, i.e. when the spectral information is translated from the coarse scale of the MS data to the fine scale of the Pan image. Experimental results carried out on Ikonos data demonstrate that the results provided by the proposed information-theoretic method are in trend with analysis performed on spatially degraded data by means of average SAM, Walds's ERGAS and the novel Q4 index based on quaternion theory and recently proposed by the authors. However, the novel method requires no reference and is therefore usable in all practical cases.

1 INTRODUCTION

Remote-sensing image fusion techniques aim at integrating the information conveyed by data acquired with different spatial and spectral resolution from satellite or aerial platforms. The main goal is photoanalysis, but also automated tasks such as features extraction and segmentation/classification have been found to benefit from fusion. A variety of image fusion techniques is devoted to merge multispectral (MS) and panchromatic (Pan) images, which exhibit complementary characteristics of spatial and spectral resolutions [1]. Injection in the re-sampled MS images of spatial details extracted from the Pan image has been found to be adequate for preserving the spectral characteristics [2]. Multiresolution analysis, based on undecimated wavelets decompositions and Laplacian pyramids, has proven itself effective to implement fusion at different resolutions [3].

Quantitative results of data fusion are provided thanks to the availability of reference originals obtained either by simulating the target sensor by means of high resolution data from an airborne platform [4], or by degrading all available data to a coarser resolution and carrying out fusion from such data. In practical cases this strategy is not feasible. The underlying assumption, however, is that fusion performances are invariant to scale changes [5]. Hence, algorithms optimized to yield best results at coarser scales, i.e. on spatially degraded data, should still be optimal when the data are considered at finer scales, as it happens in practice. This assumption may be reasonable in general, but unfortunately may not hold for very high resolution data, especially in a highly detailed urban environment, unless the spatial degradation is performed by using low-pass filters whose frequency responses match the shape of the modulation transfer functions (MTF) of the sensor [6].

As an alternative to this protocol, the problem of measuring the quality of fusion may be approached at the full spatial scale without any degradation. The spectral and spatial distortions are separately evaluated from the available data, i.e. from the original low-resolution MS bands and high resolution Pan image, and the outcomes are properly combined to yield a global quality index. According to the protocol described by Zhou [7], the spectral quality is calculated for each band as the absolute cumulative difference between the fused bands and the re-sampled input bands, while the spatial quality is measured by the correlation coefficient between the spatial details of each of the fused MS bands and those of the Pan image; such details are extracted by means of a Laplacian filter. The major drawback of this approach is that spatial and spectral distortion follow opposite trends. Spectral distortion is zero if no spatial enhancement is made. Spatial distortion is zero, or better spatial quality is maximal, if the fused MS bands are each proportional to the Pan image; in which case spectral information is lost after fusion. Furthermore, if MS bands with same resolution as the Pan image were hypothetically available, it would be found by Zhou's protocol that they are significantly distorted and hence of poor quality.

In this work, we present a global index capable of jointly measuring the spectral and spatial quality and working at the full scale, same as in Zhou's protocol [7]. The spatial and spectral distortions are separately calculated from the mutual information between the fused MS image, the source MS image, and the Pan image. A combination of the spectral and spatial distortion indices is carried out to obtain a unique quality index. The rationale is that the information inter-relationships between any couple of spectral bands and between each band and the Pan image should be unchanged after fusion. Changes in the former are responsible for spectral distortion. Changes in the latter indicate spatial distortion. The underlying assumption of inter-scale preservation of mutual information is demonstrated by the fact that true high-resolution MS data, whenever available, will exhibit spectral and spatial distortions that are both zero, within the approximations of the model, and definitely lower than those attained by any fusion method.

2 MUTUAL INFORMATION

The concept of mutual information (MI) represents a measurement of the relative entropy between two information sources, A and B, that is the measurement of information redundancy [8]. From this definition, it could be derived that the MI is maximal when the two sources coincide; in this case MI becomes equal to the autoinformation (AI), or entropy:

$$I(A;B) = \sum_A \sum_B p_{A,B}(a,b) \cdot \log \left[\frac{p_{A,B}(a,b)}{p_A(a) \cdot p_B(b)} \right] \tag{1}$$

$$I(A;A) = H(A) = -\sum_A p_A(a) \cdot \log[p_A(a)] \tag{2}$$

with $p_A(a)$, $p_B(b)$ the marginal distributions and $p_{A,B}(a,b)$ the joint distribution of A and B. MI is related to entropy by the following relationships:

$$I(A;B) = H(A) + H(B) - H(A,B) \tag{3}$$

$$I(A;B) = H(A) - H(A|B) \tag{4}$$

$$I(A;B) = H(B) - H(B|A) \tag{5}$$

with $H(A)$ and $H(B)$ being the entropy of A and B, $H(A,B)$ their joint entropy and $H(A|B)$ the conditional entropy of A given B:

$$H(B) = -\sum_B p_B(b) \cdot \log[p_B(b)] \tag{6}$$

$$H(A,B) = -\sum_A \sum_B p_{A,B}(a,b) \cdot \log[p_{A,B}(a,b)] \tag{7}$$

$$H(A|B) = -\sum_A \sum_B p_{A,B}(a,b) \cdot \log[p_{A|B}(a|b)] \tag{8}$$

While the entropy is known to measure the amount of uncertainty about the random variable, MI is the amount of information that B contains in A. Some properties of MI are summarized in Table 1.

Table 1: Some properties of MI

$I(A;B) \geq 0$
$I(A;B) = I(B;A)$
$I(A;B) \leq \min\{H(A), H(B)\}$

Several applications have shown that MI is capable of extracting the common information better than the correlation coefficient (CC) [9, 10, 11]. Whenever it is required to calculate MI on small blocks of the available data and average the results over the whole data set, algorithms based on bivariate histograms cannot be utilized because of statistical instability of the estimated probabilities. In that case, by modelling the two information sources as a locally stationary and ergodic bivariate Gaussian random process, a simple expression is derived for the local MI [8]:

$$I(A;B) = -\log\left(\sqrt{1-\rho^2_{A,B}}\right) \tag{9}$$

in which $\rho_{A,B}$ is the correlation coefficient between A and B and is calculated locally on blocks of data.

3 QUALITY INDEX WITHOUT REFERENCE

3.1 Spectral Distortion Index

A spectral distortion index can be derived from the difference of inter-band MI values calculated from the fused MS bands, indicated as $\{\hat{G}_l\}_{l=1}^L$, and from the low-resolution MS bands, re-sampled to the spatial scale of Pan, $\{\tilde{G}_l\}_{l=1}^L$. The MI terms $I(\hat{G}_l;\hat{G}_r)$ and $I(\tilde{G}_l;\hat{G}_r)$ can be grouped into two $L \times L$ matrices. The two matrices are symmetrical and the values on the main diagonal are all equal to one.

A spectral distortion index, referred to as D_λ, is calculated as

$$D_\lambda \triangleq \sqrt[p]{\frac{1}{L^2-L}\sum_{l=1}^{L}\sum_{\substack{r=1\\r\neq l}}^{L}\left|I(\hat{G}_l;\hat{G}_r) - I(\tilde{G}_l;\tilde{G}_r)\right|^p} \tag{10}$$

p being a positive integer exponent chosen to emphasize large spectral differences: for $p = 1$, all differences are equally weighted; as p increases, large components are given more relevance. The index (10) is proportional to the $p-$norm of the difference matrix, being equal to 0 if and only if the two matrices are identical. If the model reported in (9) is exploited to calculate MI, values of $I(\hat{G}_l;\hat{G}_r)$ and $I(\tilde{G}_l;\hat{G}_r)$, originated by inter-band correlation coefficients close to one, are clipped above one, so that (10) is always lower than one.

3.2 Spatial Distortion Index

A spatial distortion index is calculated as

$$D_s \triangleq \sqrt[q]{\frac{1}{L}\sum_{l=1}^{L}\left|I(\hat{G}_l;P) - I(\tilde{G}_l;\tilde{P})\right|^q} \qquad (11)$$

in which P is the Pan image and $\tilde{P}$ a low-pass version of the Pan image obtained by filtering the Pan image with a low-pass filter having normalised frequency cutoff at the resolution ratio between MS and Pan. Analogously, D_s is proportional to the $q-$norm of the difference vector, where q is chosen so as to emphasize higher difference values. The index D_s attains its minimum, (equal to zero) when the two MI vectors are identical. Analogously to (10), also (11) is upper bounded by one because the clipping above one of MI values between MS and Pan is enabled. However clipping occurs very seldom, because the local correlation between MS and Pan is moderate.

3.3 Jointly Spectral and Spatial Quality Index

The use of two separate indices, may be not sufficient to establish the ranking of performances of fusion methods. In fact, D_λ and D_s respectively measure changes in spectral behaviour occurring between the re-sampled original and the fused images and discrepancies in spatial details originated by fusion.

To trade off the above trends, let us introduce a single index, namely QNR, i.e. *Quality with No Reference*, which is the product the one's complements of the spatial and spectral distortion indices, each raised to a real-valued exponents that separately attribute the relevance of spectral and spatial distortions to the overall quality and jointly determine the non-linearity of response in the interval $[0,1]$, same as a γ to achieve a better discriminations of the fusion results compared:

$$\text{QNR} \triangleq (1 - D_\lambda)^\alpha \cdot (1 - D_s)^\beta. \qquad (12)$$

Thus, the highest value of QNR is one and is obtained when the spectral and spatial distortions are both zero. The main advantage of the proposed index is that, thanks to the lack of a reference data set, quality can be assessed at the full scale of Pan. Experimental results aimed at validating the proposed index are shown in the next section.

4 EXPERIMENTAL RESULTS

The proposed quality index has been assessed on very high-resolution image data collected by the Ikonos space-borne MS scanner on the city of Toulouse, France. The four MS bands of Ikonos span the visible and NIR wavelengths and are non-overlapped, with the exception of B1 and B2: B1=440÷530 nm, B2=520÷600 nm, B3=630÷700 nm and B4=760÷850 nm. The bandwidth of Pan embraces the interval 450÷950 nm. The data set has been radiometrically calibrated from digital counts and geo-coded to 4 m (MS) and 1 m (Pan) GSD. A square region of 4.2 km^2 was analysed. The original Pan image is of size 2048×2048 and the original MS image of size 512×512.

The following fusion methods have been compared:

- GLP-based method with context-based decision model (CBD) [3, 6];
- Generalised IHS-based method with injection model based on genetic algorithms (GIHS-GA) [12];
- Gram-Schmidt spectral sharpening method (GS) [13], as implemented in ENVI [14];
- Intensity-Hue-Saturation method with spectral adjustment (Tu-IHS) [15];

The first experiment aims at demonstrating that QNR, which does not require reference original, is in accordance with other score indices that require reference originals, like:

- Spectral Angle Mapper (SAM) denotes the absolute value of the spectral angle between two vectors, $\mathbf{v}$ and $\hat{\mathbf{v}}$,

  $$\text{SAM} = \arccos\left(\frac{< v, \hat{v} >}{\|v\|_2 \cdot \|\hat{v}\|_2}\right). \qquad (13)$$

  SAM equal to zero denote absence of spectral distortion, but possible radiometric distortion. SAM is usually averaged over the whole image to yield a global distortion index.

- Vectorial Root Mean Square Error (VRMSE) is an index that measures the overall radiometric distortion

  $$\text{VRMSE} = \sqrt{\frac{1}{L}\sum_{l=0}^{L-1}(\text{RMSE}(l))^2} \qquad (14)$$

- ERGAS, which means relative dimensionless global error in synthesis, [16, 17], is given by

  $$\text{ERGAS} = 100\frac{d_h}{d_l}\sqrt{\frac{1}{L}\sum_{l=0}^{L-1}\left(\frac{\text{RMSE}\,(l)}{\mu(l)}\right)^2} \qquad (15)$$

  where d_h/d_l is the ratio between the pixel sizes of the Pan and MS, e.g. 1/4 for QuickBird and Ikonos data, $\mu(l)$ is the mean of the lth band. This score index measures a distortion and thus must be as small as possible.

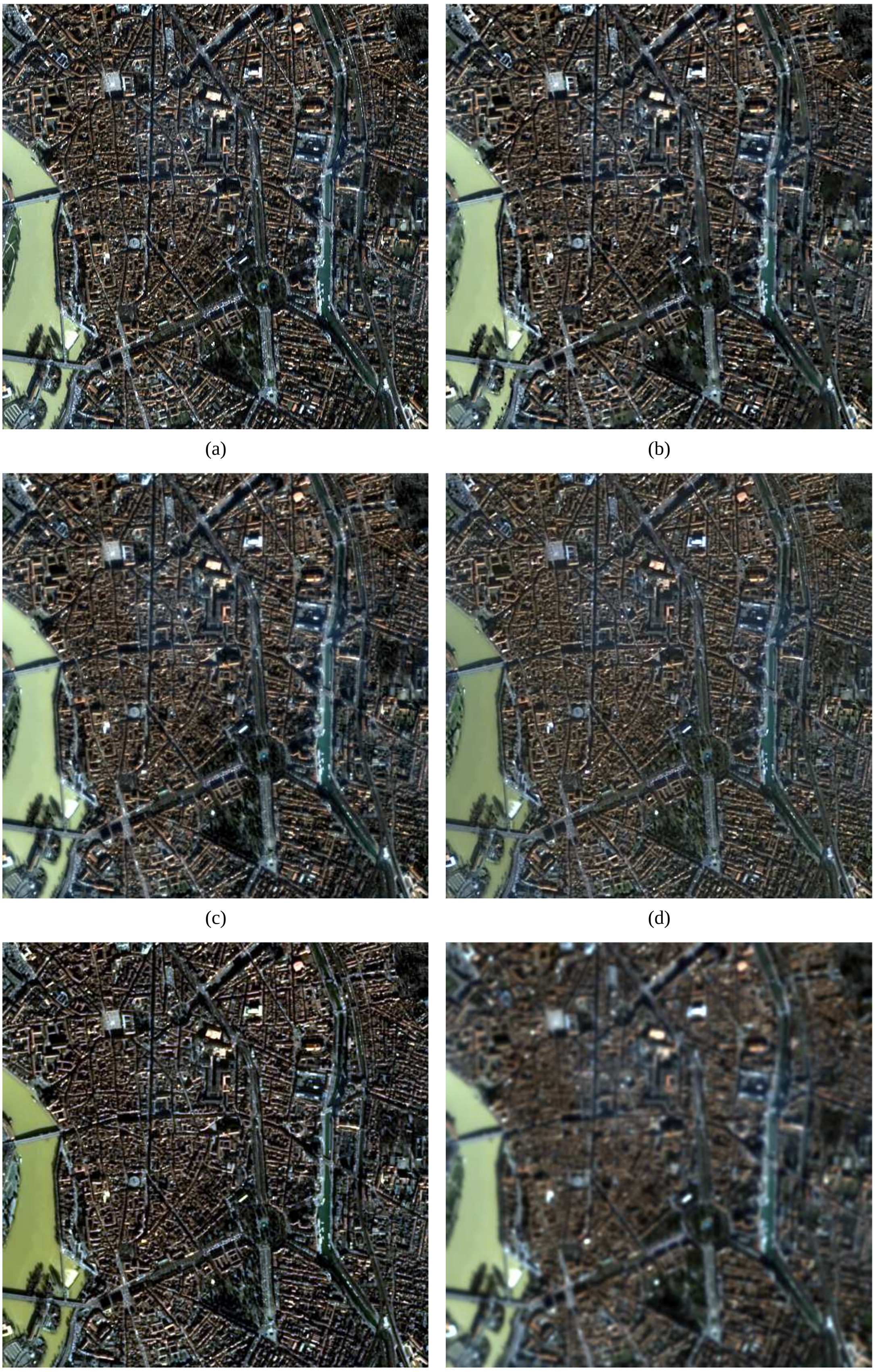

Figure 1: True color composite of 3-2-1 bands of (a): true MS bands, (b): CBD fusion, (c): GIHS-GA fusion, (d): GS fusion, (e): Tu-IHS fusion, (f): Re-sampled MS bands.

- The quality index Q4 [18] is a generalisation to 4-band images of the Q index [19], which can be applied only to monochrome images. Q4 is obtained through the use of CC between hypercomplex numbers, or *quaternions*, representing spectral pixel vectors. Q4 is made of three different factors:

$$Q4 = \frac{|\sigma_{z_1 z_2}|}{\sigma_{z_1} \cdot \sigma_{z_2}} \cdot \frac{2\sigma_{z_1} \cdot \sigma_{z_2}}{\sigma_{z_1}^2 + \sigma_{z_2}^2} \cdot \frac{2|\bar{\mathbf{z}}_1| \cdot |\bar{\mathbf{z}}_2|}{|\bar{\mathbf{z}}_1|^2 + |\bar{\mathbf{z}}_2|^2} \quad (16)$$

  the first is the modulus of the hypercomplex CC between the two spectral pixel vectors and is sensitive both to loss of correlation and to spectral distortion between the two MS data sets. The second and third terms respectively measure contrast changes and mean bias on all bands simultaneously. The modulus of the hypercomplex CC measures the alignment of spectral vectors. Therefore, its low value may detect when radiometric distortion is accompanied by spectral distortion. Thus, both radiometric and spectral distortions may be encapsulated in a unique parameter. All statistics are calculated as averages on $N \times N$ blocks, either $N = 16$ or $N = 32$. Eventually, Q4 is averaged over the whole image to yield the *global* score index. The highest value of Q4, attained if and only if the test MS image is equal to the reference, is one; the lowest value is zero.

Table 2: Global quality indices of fused Ikonos data. QNR is calculated without reference, unlike the other indices.

	SAM	VRMSE	ERGAS	Q4	QNR
REF	0	0	0	1	0.928
CBD	3.214	18.31	3.041	0.899	0.654
GIHS-GA	3.335	20.30	3.391	0.882	0.633
GS	4.786	27.93	4.716	0.780	0.649
Tu-IHS	4.366	23.17	3.788	0.873	0.548
EXP	4.968	36.20	6.070	0.535	0.601

Table 3: QNR values calculated at degraded (4 m) and full resolution (1 m) on the same geographical area.

QNR	4m	1m
CBD	0.654	0.745
GIHS-GA	0.633	0.726
GS	0.649	0.656
Tu-IHS	0.548	0.697
EXP	0.601	0.670

To this purpose, according to the protocol proposed in [5], the data sets have been spatially degraded by four, and statistics have been calculated, for all indices except QNR, between fused and original data. Fusion results are shown in Fig. 1. From the numerical values reported

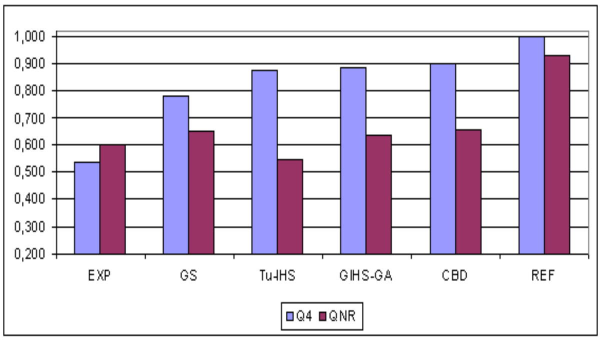

Figure 2: QNR and Q4 indices calculated at degraded scale (4 m).

in Table 2, two considerations can be made. The QNR value of the 4 m reference original (REF) is close to one and far greater than those of any other method. QNR values, calculated without reference original, are substantially in accordance with the other scores, which require reference originals. Discrepancies are due to the fact that the exponents α and β in (12) are taken both equal to one. Thus spectral and spatial quality are given the same importance. Instead, SAM measures mainly spectral distortion, VRMSE is little sensitive to spectral distortion, while in ERGAS and Q4 it is impossible to quantify the sensitiveness to spectral and spatial distortion separately. This explain why the re-sampled image without enhancement (EXP), which has spectral distortion practically zero exhibits better QNR than other methods providing a spatial enhancement. However, for all five indices, CBD attain global scores better than those of the other method, followed by GIHS-GA. Fig. 2 shows the trends of QNR and Q4 indices varying with fusion methods applied at degraded resolution.

The second experiment aims at comparing the numerical values of QNR across scales, i.e. on both fused images obtained from degraded originals and fused images at the full scale of Pan. Although values of QNR depend on the scale, Table 3 evidences that performances of fusion methods are roughly similar for the two scales. A notable exception is that the performance ranks of GS and Tu-IHS are swapped from one scale to the other.

5 CONCLUSIONS

A new quality index of pan-sharpened MS images has been developed, based on the evidence that the mutual information relationships between couples of bands and between each band and the Pan image are unchanged from one scale to another. Thus, the original MS and Pan data can be used to measure the spectral and spatial distortion, without resorting to spatial degradation of the data-set to a coarse scale. Experimental results, carried out on Ikonos data by means of a number of fusion methods, demonstrate that the results provided by the proposed information-theoretic method are in trend with analysis performed on spatially degraded data, as well as that performances of fusion methods actually depend on the spatial scale on which fusion is accomplished. Future developments are focussed on finding another tool to measure inter-band relationships, more fitting and flexible than mutual information.

ACKNOWLEDGEMENT

The authors wish to thank Massimo Selva of IFAC-CNR for his valuable support in computer simulations of fusion methods. This work has been supported in part by MIUR under Cofin 2005.

References

[1] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, “A comparative analysis of image fusion methods,” *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 6, pp. 1391–1402, June 2005.

[2] P. S. Chavez Jr, S. C. Sides, and J. A. Anderson, “Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic,” *Photogramm. Eng. Remote Sens.*, vol. 57, no. 3, pp. 295–303, 1991.

[3] B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli, “Context-driven fusion of high spatial and spectral resolution data based on oversampled multiresolution analysis,” *IEEE Trans. Geosci. Remote Sensing*, vol. 40, no. 10, pp. 2300–2312, Oct. 2002.

[4] F. Laporterie-Déjean, H. de Boissezon, G. Flouzat, and M.-J. Lefèvre-Fonollosa, “Thematic and statistical evaluations of five panchromatic/multispectral fusion methods on simulated PLEIADES-HR images,” *Inform. Fusion*, vol. 6, no. 3, pp. 193–212, Sep. 2005.

[5] L. Wald, T. Ranchin, and M. Mangolini, “Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images,” *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691–699, 1997.

[6] B. Aiazzi, L. Alparone, S. Baronti, A. Garzelli, and M. Selva, “MTF-tailored multiscale fusion of high-resolution MS and Pan imagery,” *Photogramm. Eng. Remote Sens.*, vol. 72, no. 5, pp. 591–596, May 2006.

[7] J. Zhou, D. L. Civco, and J. A. Silander, “A wavelet transform method to merge Landsat TM and SPOT panchromatic data,” *Int. J. Remote Sensing*, vol. 19, no. 4, pp. 743–757, 1998.

[8] R. E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley, Reading, MA, 1987.

[9] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Seutens, “Multimodality image registration by maximization of mutual information,” *IEEE Trans. Medical Imaging*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[10] P. Thevenaz and M. Unser, “Optimization of mutual information for multiresolution image registration,” *IEEE Trans. Image Processing*, vol. 9, no. 12, pp. 2083–2099, Dec. 2003.

[11] A. A. Cole-Rodes, K. L. Johnson, J. LeMoigne, and I. Zavorin, “Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient,” *IEEE Trans. Geosci. Remote Sensing*, vol. 12, no. 12, pp. 1495–1511, Dec. 2003.

[12] A. Garzelli and F. Nencini, “Fusion of panchromatic and multispectral images by genetic algorithms,” in *Proc. IEEE Int. Conf. on Geosci. Remote Sensing*, July 2006.

[13] C. A. Laben and B. V. Brower, “Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening,” Tech. Rep. US Patent # 6,011,875, Eastman Kodak Company, 2000.

[14] ENVI® Version 4.1 User Manual, Research Systems Inc., 2004.

[15] T.-M. Tu, P. S. Huang, C.-L. Hung, and C.-P. Chang, “A fast Intensity-Hue-Saturation fusion technique with spectral adjustment for IKONOS imagery,” *IEEE Geosci. Remote Sensing Lett.*, vol. 1, no. 4, pp. 309–312, Oct. 2004.

[16] T. Ranchin and L. Wald, “Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation,” *Photogramm. Eng. Remote Sens.*, vol. 66, no. 1, pp. 49–61, Jan. 2000.

[17] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, “Image fusion - the ARSIS concept and some successful implementation schemes,” *ISPRS J. Photogramm. Remote Sensing*, vol. 58, no. 1-2, pp. 4–18, June 2003.

[18] L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, “A global quality measurement of Pan-sharpened multispectral imagery,” *IEEE Geosci. Remote Sensing Lett.*, vol. 1, no. 4, pp. 313–317, Oct. 2004.

[19] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Processing Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.